WHEN SAMPLE QUANTITATION LIMITS ARE NEAR THE PROJECT DECISION POINT: AN APPROACH TO PLANNING FOR ACCEPTABLE LEVELS OF UNCERTAINTY IN THE RESULT

Bill Wright[1,2], Montgomery Watson, Seattle Industrial/Federal Operations Program
John Wakeman[3], US Army Corps of Engineers, Seattle District
Travis Shaw[3], US Army Corps of Engineers, Seattle District

Abstract

This paper describes a method (distribution analysis) which was used to deal with a problem arising during project planning for an environmental investigation. The problem: limitations of analytical methods can cause problems for data quality objectives (DQOs). DQOs are conditions for generating usable data of sufficient quality for a project decision, including project uncertainties. In the present case, environmental risk-based concentrations (RBCs) have been calculated to be extremely low relative to sample quantitation limits, so that confirming the presence or absence of toxicants with a degree of certainty was difficult, in turn setting up a situation where risk could be "hidden" in the results in the form of non-detected analytes. DQOs for uncertainty (specifically, power = 0.80) established during this project resulted in a well-documented, successful investigation. Although the situation discussed here was unusual, the problem addressed is general, and the method of distribution analysis is readily available for broader use.

The Camp Bullis Landfill 8 project targeted chemical warfare materiel (CWM) compounds (Mustard and Lewisite and their break-down products) suspected to occur in site groundwater north of San Antonio, Texas. Recently, the U.S. Army Center for Health Promotion and Preventative Medicine (the office of the Army Surgeon General) developed oral reference doses for several CWM compounds which predict toxicity at levels close to drinking-water method detection limits (MDLs) and sample quantitation limits (SQLs). Available analytical techniques are very limited and tend to be proprietary, both within the Army and in private laboratories; hence they are difficult to improve. During the project planning phase it was also realized that time and budget constraints of this project did not allow for the refinement of such analytical techniques to increase their sensitivity.

During the planning process for this project, distributional (or probabilistic) analysis was used to incorporate available information from laboratory MDL studies, in order to evaluate prospective false-negative error rates. Such consideration notably improved our view of false negative error. DQO attainment was evaluated again after the data acquisition phase, and this confirmed the planning stage predictions of acceptable limits of error. The analytical methods were deemed adequate to determine the lack of risk to the public from ingestion of groundwater at the site which could have been contaminated with CWM.

[1] Montgomery Watson, Inc.; 2375 130th Ave. NE, Suite 200, Bellevue, Washington 98005-1758; 425-881-1100 phone; 425-867-1970 fax; william.e.wright@us.mw.com
[2] Author to whom correspondence should be addressed.
[3] Seattle District; CENWS-ED-TB-ET; POB 3755, Seattle, Washington 98124; 206-764-3430 phone; 206-764-6795 fax; john.s.wakeman@usace.army.mil and travis.c.shaw@usace.army.mil.

.

Introduction

Camp Bullis is a 27,900-acre Army training camp that was established north of San Antonio, Texas in 1890. It is used by all branches of the military for combat, parachute, medical, and ground defense training. Fort Sam Houston, which oversees Camp Bullis, has implemented a closure program for 21 landfills at the camp. In 1994, officials discovered agents on the eroding landfill surface OF Landfill 8, a 9-acre facility, 1935-vintage Navy chemical agent identification sets used in training, and both broken and intact containers which were used for holding agents. The landfill was fenced off until CWM characterization in drinking water supplies could be addressed. Landfill 8 is thought to be near recharge areas for the Edwards Aquifer which provides drinking water supply for greater San Antonio.

In the fall of 1996, Fort Sam Houston retained the Seattle District to investigate CWM at Landfill 8. The Seattle District retained Montgomery Watson to assist on the project team. The objective was to plan and implement a groundwater investigation of Landfill 8 so as to characterize the nature and extent of any CWM contamination. A constrained project budget and schedule allowed the installation of only three monitoring wells to be sampled, in one round.

Corps of Engineers' technical documents which guide planning for hazardous, toxic, and radioactive waste investigations emphasize the importance of the DQO development process. The key project decision was whether to remediate Landfill 8; the primary site characterization decision needed to support the ultimate decision on remediation was whether a significant release of CWM to groundwater could be confirmed.

Methods

The primary site characterization objective was formulated as the following null and alternate hypotheses:

> H_0—there is no significant release of CWM to groundwater; and,
> H_a—there is a significant release of CWM to groundwater.

While documenting the conceptual site model, it became obvious to the project planning team that detection of CWM and breakdown products in groundwater was rather unlikely—the CWM has been found on the landfill surface in the form of small ampules and bottles. These bottles are seldom intact, and each contained very small quantities (a few mL) of agents which are known to be unstable when released to the environment. The challenge within the planning framework was to design the investigation such that a single sample from a monitoring well with a concentration reported nondetected, could be regarded as proof of the absence of a significant release of CWM to groundwater. At the time of the investigation, the Army had just issued chronic oral reference doses (chronic RfD's) for CWM and Texas regulates on the basis of risk. This study is, to our

knowledge, the Army's first CWM site investigation performed in light of these RfDs. Previous studies have used the Army's long-established, occupational drinking-water standards (based on an acute combat exposure scenario) as benchmarks for risk. The chronic RfDs yielded substantially lower risk-based concentrations (RBCs) in groundwater than the Army drinking-water standard. Although, as discussed below, it was expected that the site values would be close to zero and probably below the Method Detection Limit, MDL[1], the project decision was whether quantified values of CWM were present (that is, whether detection occurred at or above the Sample Quantitation Limit, SQL). This was because the Army would regard, by policy, confirmed presence of CWM compounds such as Mustard or Lewisite in groundwater as sufficient evidence to classify the medium as hazardous. It is not clear whether an estimated value (J, between the MDL and SQL) would be regarded as confirmed presence. For this reason, the following discussion references the MDL for the planning phase when the SQL was unknown, and the SQL for the post-project evaluation.

The hypothesis testing challenge was the need to control the Type II error rate, β. See Table 1 on the following page. In a Type II error, one accepts a false null hypothesis, in which case a decision would be made that no risk exists, when risk does in fact exist. In hypothesis testing, scientists usually only control the Type I or false-positive error rate, α. This H_0 and H_a are usually formulated in such a way that H_0 is the simpler statement of the two. H_0 is thus usually the working hypothesis unless it can be rejected at a specified level of Type I error (typically, $\alpha = 0.05$). The vast majority of statistics textbooks address control of α or Type I error only, and do not discuss methods for controlling β. For this reason, one seldom speaks of either rejecting H_a or accepting H_0; H_0 is either retained or rejected, and, if rejected, H_a is then accepted as the working hypothesis.

---

[1] The method detection limit (MDL) is defined as the minimum concentration of a substance that can be measured and reported, with 99% confidence, that the analyte concentration is greater than zero and is determined from analysis of a sample in a given matrix containing the analyte. Most laboratories perform MDL studies on an annual basis, at a minimum, in an interference-free matrix; typically reagent water for water methods. The Method Quantitation Limit (MQL) is defined as the lowest calibration standard and should be no lower ten times the standard deviation as determined from the MDL study.

The Sample Quantitation Limit (SQL) is a laboratory-determined reporting limit which may be based on a project-specific reporting limit, a regulatory action level, or the laboratory's MQL. Often, in the absence of project-specific reporting limits, the SQL is set at the value of the low standard used during initial calibration. The low standard is typically at or below the project-specific action levels or regulatory levels. Analytical values can be reported by the laboratory below the low standard, but should then be reported as estimated values based upon the analyst's judgment.

Table 1: Hypothesis Testing

		Hypothesis	
		Actually True	Actually False
Decision	Accept	**1-α**	β
	Reject	α	**1-β**

When β is specified in a DQO, and the DQO is met in a hypothesis test, one can reject H_a and thus accept H_0. Whereas α is typically specified at 0.05, β, when defined by the planners, is usually specified at 0.20 in Superfund programs. The complement of α, 1-α, specifies the level of confidence of the hypothesis test; and the complement of β, 1- β, specifies the power of the test, i.e., the ability to correctly reject a false H_0. The conventional confidence level 1-α is 95%, and the conventional power level 1- β is 80%.

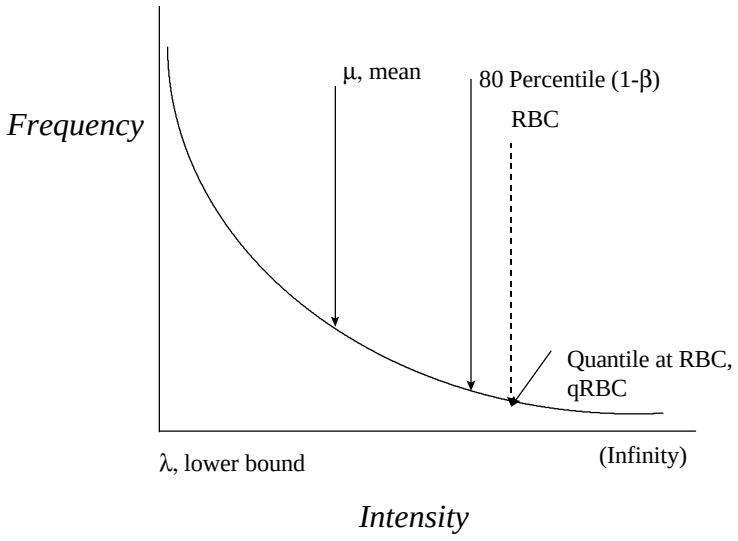
The project planning team set out to design a limited investigation that could reject H_a, "there is a significant release of CWM to groundwater," at a power level of 80%. However, most limited sampling designs have extremely low power. Everything else being constant, the only way to increase the power of a hypothesis test is to either increase sample size or decrease the level of confidence. Given project constraints, the first was not possible. In fact, with minimal replicate sampling (a consequence of project budget constraints), an objective, classical statistical approach to hypothesis testing was probably not possible; therefore, it was also not possible to decrease the level of confidence.

The team decided to use an information theory approach to controlling β. A simple probabilistic model, a probability distribution, was developed that acknowledged, incorporated, and considered the effects of uncertainties occurring during sample analysis. For each CWM analyte at each well (with the exception of one well which had a field duplicate), a single, validated, somewhat uncertain, analytical result would be available along with a limited amount of other information. For example, due to physical constraints, the uncertain concentration of each CWM analyte could be no less than 0 and no greater than about 1 billion μg/L (a 100% solution). As the single analytical result for each well was expected to be less than an SQL (0.02 and 0.93 μg/L for Mustard and

Lewisite, respectively), the upper bound could be regarded, for all practical purposes, as infinity.

Information theory was utilized in this investigation. Shannon and Weaver (1949) state that maximum entropy of information is a measure of maximum uncertainty. Given any set of information, there are a infinite number of probability distributions that can be fitted to the information available. There is only one distribution, however, that is based on only the information at hand—the one that maximizes the informational entropy or uncertainty within the context of the available knowledge constraints. Such a subjective probability distribution (objectively derived) is known as a state-of-knowledge distribution. This is conservative in terms of predicting environmental consequences when placed in the context of RBCs. Shannon and Weaver showed that when the knowledge constraints consist only of a lower bound, λ, of 0 and a mean, μ, the most uncertain distribution one can define mathematically is the exponential distribution, called $\mathbf{e}(\mu)$.

Figure 1. Example of Exponential Distribution, Illustrating Terms in Text



The project team believed that they could define a defensible state-of-knowledge distribution on the basis of an estimate of the mean concentration of a CWM analyte in a given monitoring well. The problem remained, however, that the expected result of "less than the SQL" would not directly yield an estimate of the mean. Fortunately, an exponential distribution is a single-parameter distribution—once the mean is defined the entire distribution is defined. Conversely, once any other parameter is defined, such as a quantile of the distribution, the mean and thus the entire distribution can be defined (Montgomery Watson and Envirochem, 1995). The procedure sought to define the distribution in the best manner possible, and to disclose where the RBC fell on this distribution, its quantile. From this, it is possible to estimate as well a mean

environmental concentration for the highest occurrence of the compound of concern in the medium.

By way of illustration, Montgomery Watson previously applied this approach to a problem faced by the British Columbia Ministry of the Environment. The Ministry had a site contaminated with arsenic for which they had obtained 31 air samples. The reported arsenic concentration in each sample was reported as being less than the SQL, which was, unfortunately, well above the Ministry's risk-based concentration. The Ministry's initial position was that these data were inadequate and that further sampling, using an alternative analytical technique with a lower MDL, would be needed to quantify the atmospheric arsenic concentrations at the site. With an understanding that we often know more than what we initially appear to know and of subjective probability assessment, Montgomery Watson was able to defensibly quantify the arsenic concentrations so as to avoid the costs of additional sampling and analysis and the associated delay in the project schedule. The approach used is outlined below.

It was assumed that the arsenic concentrations could vary anywhere from 0 to virtually infinity. If one has n samples ranked from lowest to highest, where r is the rank value for each sample, and r varies from 1 to n, the r^{th} sample provides an estimate of the (r-0.5)/n quantile of the underlying probability distribution of concentration values. Therefore, the highest concentration sample, at which r equals n, estimates the (n-0.5)/n quantile. Of course, if all n samples are reported as "< SQL," then we don't have an estimate of the (n-0.5)/n quantile. However, the SQL can be used as a conservative surrogate of the unknown, but lower than SQL value of n^{th} value.

In the BC Environment example, n was 31 and the SQL was 0.1 $\mu g/m^3$. Therefore, using equations shown below, 0.1 $\mu g/m^3$ was conservatively considered to be on the 0.984 quantile of the distribution. The relationship between the mean, μ, of an exponential distribution and a quantile, q, is:

$$\mu = \frac{-x_q}{\ln(1-q)} \qquad \text{Equation 1}$$

where x_q is the value at the q quantile. With an x_q of 0.1 $\mu g/m^3$ and a q of 0.984, μ is 0.0242 $\mu g/m^3$. Thus, BC Environment could thus represent the uncertain background arsenic concentration in air as **e**(0.0242) $\mu g/m^3$, which was below their RBC value, and further sampling was determined to be unnecessary.

The remainder of this paper will focus on Mustard and Lewisite as key analytes for the Landfill 8 project's DQO planning and evaluation. The distribution approach was chosen to characterize the CWM data from the Camp Bullis groundwater investigation. If the single result, per analyte and per well, was reported as "< SQL," x_q would be set at the SQL and q would be set at 0.5.

The Army Surgeon General's recently published RfDs for Mustard and Lewisite, their corresponding risk-based concentrations (RBCs, or concentrations corresponding to a hazard quotient of 1.0), the laboratory's MDL, and the reported SQL are shown in Table 2.

Table 2. Toxicity and Analytical Benchmarks for Mustard and Lewisite.

CWM	*RfD, μg/(kg·d)*	*RBC, μg/L*	*MDL, μg/L*	*SQL, μg/L*
Mustard	0.007	0.26	0.03[2]	0.02
Lewisite	0.1	3.65	0.03	0.93

A commercial analytical laboratory was contracted for this project because most of the available methods are proprietary. Also, as noted, there were constraints which did not permit a substantial research and development effort to reduce the MDLs. The commercial laboratory provided its MDL studies for CWM analytes using its own proprietary methods, and in fact reran them with better results than in the planning phase just prior to the investigation analyses. The project work plan evaluated the commercial laboratory MDLs to ensure that β error rates were sufficiently low to allow for the attainment of DQOs in the event that the CWM analyte concentrations would be reported, as anticipated, as "< MDL" (see Table 3). The plan concluded that it was feasible to expect a 1- β of $\geq$80%.

[2] This value refers to the planning stage MDL.. The lab repeated their MDL studies before the investigation was mounted, and improved their performance. Usually, the SQL is about three times the MDL, but it can be higher.

Table 3. Planning Estimates for Mustard and Lewisite, Using a 1,000 Trial Monte Carlo Simulation.

CWM	μ	q_{RBC} *or* 1- β *without taking into account spike recoveries*	*MDL Study Spike Range*	*MDL Study Spike Mean* +/- σ	q_{RBC} *or* 1- β *taking into account spike recoveries*	*RBC Value* (u*g/L)*
Mustard	0.03	99.99	0-100%	89 +/- 7.1	93	0.26
Lewisite	0.03	98	0-117%	54 +/- 21	90	3.65

An additional factor in the DQO development was the representativeness of the groundwater samples; *i.e.*, for the project to be successful, the wells had to be installed in locations such that any release from the landfill to the groundwater would be intercepted. This could not be evaluated *a priori* as no definitive information about the direction of groundwater flow would be available until after the wells were installed.

Results

As predicted by the project planning team, no CWM analytes were detected, in any of the wells sampled, during the single round of groundwater sampling and analysis. All DQOs for quality assurance/quality control specified in the quality assurance project plan were met for Mustard and Lewisite.

For Mustard, the results of all samples (one per well) of < 0.02 μg/L equates to **e**(0.029) μg/L. For Lewisite, < 0.93 μg/L (a disappointing result, given that it is 31 times the MDL) equates to **e**(1.3) μg/L. (Table 4.) The value of β was determined simply by determining where the RBC lies on the corresponding distribution:

$$\beta = 1 - q_{RBC} \qquad \text{Equation 2}$$

where q_{RBC} is the quantile corresponding to the RBC. This value is calculated by rearranging Equation 1 (and substituting RBC for x_q), as follows:

$$q_{RBC} = 1 - e^{(-RBC/\mu)}. \qquad \text{Equation 3}$$

As expected, no CWM detections occurred. The values for q_{RBC} and β for Mustard and Lewisite are therefore:

Table 4. Mustard and Lewisite Results Compared to RBCs, with Associated False Negative Error Rates

CWM	q_{RBC}	μ	β	RBC
Mustard	0.9999	0.029	0.0001	0.26
Lewisite	0.94	1.3	0.06	3.65

With regard to gauging the representativeness of the well samples, two of the wells installed were found to contain trichloroethene, likely originating from the landfill. Therefore, the project team concluded that the wells were positioned such that any significant release of CWM would also likely have been detected.

Discussion

Some laboratories increase their MDLs and SQLs arbitrarily to the lowest initial calibration standard used. Since CWM (in particular) calibration ranges are not specified by promulgated methods, this is a key thing to watch while planning and executing HTRW projects. This comment may be generalized to other methods, and may actually result in improved performance.

It is crucial for data generators and data users to work together to specify analytical methods that will adequately express the uncertainty of the result. This project demonstrated that numerical DQOs for uncertainty can be established during project planning so that the a winning investigation may result. As indicated by the Mustard and Lewisite results discussed in this paper, the DQO performance goal of $\beta \leq 0.2$ was met.

The ideal situation exists when the MDL is far less than the concentration representing the decision threshold. As the MDL approaches the decision threshold, decisions become more difficult. The conservative equations presented above can be used to demonstrate that, with only a single round of sampling, the SQL can not exceed 43% of the decision threshold if the default β of 0.2 is not to be exceeded. The methods used here are admittedly conservative; the following paragraph proposes a refinement which was, however, not used for this project.

Note for future development.

When the decision threshold is at the MDL, there is a nominal 1-β of 0.5. In this paper, this has been reflected in the conservative assumption that the MDL (and SQL) is the median of the distribution. However, it might be more realistic to say that it is somewhere between 0 and the MDL. Therefore, the median can be defined, applying information theory, as the uniform distribution U(0, MDL). The resulting series of all possible exponential distributions generates a distribution that is somewhat exponential in form but more skewed than exponential. This distribution would result in the MDL at

approximately the 75th percentile. In other words, it could be said that, if the RBC is at the MDL and one sample is reported as < MDL, the there is only a 25% chance of the compound being present above the RBC. Had this been applied at Camp Bullis, the power of design would have been substantially higher; but this was not needed.

Literature Cited

Montgomery Watson and Envirochem, 1995, *Stochastic Risk Assessment Training Manual*, British Columbia Ministry of the Environment, Lands and Parks and Ministry Responsible for Multiculturalism and Human Rights, Victoria, British Columbia.

Shannon, C. E., and W. Weaver, 1949, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, Illinois.